

ВІДГУК

Офіційного опонента на дисертаційну роботу

Здора Костянтина Андрійовича

На тему: «Моделі та програмні засоби підвищення швидкодії визначення

відеоатрибутів за допомогою розбиття на сцени»

представлену на здобуття наукового ступеня доктора філософії

в галузі знань 12 – Інформаційні технології

за спеціальністю 121 – Інженерія програмного забезпечення

Актуальність теми дисертації.

У сучасному світі обсяг відеоконтенту стрімко зростає, що зумовлює нагальну потребу в ефективних методах його аналізу. Визначення відеоатрибутів та розбиття відео на сцени є ключовими задачами, що дозволяють структурувати та осмислювати великі масиви відеоданих, що має важливе значення для індексації, пошуку та узагальнення інформації. Традиційні методи аналізу відео часто обмежені у своїй точності та швидкодії, особливо при роботі з динамічним контентом та складними переходами. Впровадження алгоритмів машинного навчання, зокрема глибоких нейронних мереж та візуальних трансформерів, відкриває нові перспективи для підвищення якості аналізу відео. Однак, висока обчислювальна складність таких моделей створює перешкоди для їх широкого застосування, особливо в системах реального часу та на пристроях з обмеженими ресурсами. Тому розробка моделей та програмних засобів, спрямованих на підвищення швидкодії визначення відеоатрибутів шляхом розбиття на сцени та оптимізації моделей, є надзвичайно актуальною темою.

Оцінка обґрунтованості наукових результатів дисертації, їх достовірності і новизни.

Головний науковий результат роботи полягає у розробці здобувачем моделей та програмних засобів для підвищення точності та швидкодії розбиття відео на сцени та визначення відеоатрибутів. Отримані результати мають наступну наукову новизну:

Вперше розроблено архітектуру розподіленого програмного забезпечення для визначення атрибутів на відео, яка оперує відеопотоками для їхнього розбиття на

плани та сцени, що дозволило збільшити швидкість аналізу відеоконтенту мінімум в 2.5-3 рази.

Вперше розроблено метод для виявлення переходів планів у відеоконтенті на основі поєднання математичних підходів та рекурентних нейронних мереж, який, на відміну від існуючих, швидко та ефективно виділяє просторові та часові ознаки кадрів, що дозволило збільшити точність влучання та F1-оцінку.

Вперше розроблено метод виявлення зміни сцени для відеоконтенту з використанням нейронної мережі на основі архітектури візуального трансформера для відео із застосуванням методу прунінгу перед навчанням, що, на відміну від існуючих, виділяє контекстуальні особливості сцен, дозволивши збільшити F1-оцінку на 5.1% та пришвидшити час виконання на 10%.

Набув подальшого розвитку метод прунінгу перед навчанням для моделей архітектури візуальних трансформерів для відео, який, на відміну від існуючих, враховує важливість механізму «уваги» та дозволяє пришвидшити час виконання моделі на 10%.

Достовірність отриманих в роботі наукових результатів та висновків забезпечується використанням сучасних методів дослідження, обґрунтованим застосуванням математичного апарату та підтверджується експериментально. Результати узгоджуються з сучасними теоретичними положеннями в галузі машинного навчання та аналізу відео.

Отже, в дисертаційній роботі поставлене наукове завдання виконане повністю, здобувач повною мірою оволодів методологією наукової діяльності.

Оцінка змісту дисертації, її завершеність та дотримання принципів академічної доброчесності

За своїм змістом, дисертація Здора К.А. повністю відповідає Стандарту вищої освіти зі спеціальності 121 – Інженерія програмного забезпечення та напрямкам дослідження відповідно до освітньої програми «Інженерія програмного забезпечення».

Дисертаційна робота є завершеною науковою працею і свідчить про наявність особистого внеску здобувача у розвиток методів аналізу відеоконтенту.

Розглянувши дисертаційну роботу, можна зробити висновок, що вона є результатом самостійних досліджень здобувача. Використані ідеї, результати і тексти інших авторів мають належні посилання на відповідне джерело. Порухень принципів академічної доброчесності не виявлено.

Мова та стиль викладення результатів

Дисертація написана українською мовою, є добре структурованою. Оформлення роботи послідовне та чітко підкреслює найважливіші отримані результати. Робота складається зі вступу, чотирьох розділів, висновків, списку використаних джерел та додатків. Загальний обсяг дисертації становить 178 сторінок, з яких 154 сторінки основного тексту, 3 додатки на 8 сторінках, та містить 24 рисунки, 34 формули, 9 таблиць.

У вступі розглядається актуальність дисертаційного дослідження, зв'язок роботи з науковими програмами, планами та темами. Формулюється мета, завдання, об'єкт та предмет дослідження, наукова та практична новизна отриманих результатів. Приводяться відомості про особистий внесок здобувача та апробацію результатів дисертації.

Перший розділ аналізує існуючі підходи до визначення відеоатрибутів, розглядаючи традиційні математичні алгоритми та глибокі нейронні мережі, зокрема ефективні візуальні трансформери для аналізу просторово-часових залежностей. Також досліджено методи оптимізації, такі як прунінг, квантування, дистиляція знань, для зменшення складності та пришвидшення моделей на обмежених ресурсах, з акцентом на прунінг на етапі навчання. Аналіз виявив недоліки існуючих підходів, а саме низьку точність визначення планів для сучасного контенту, нездатність враховувати контекст сцен та відсутність методів прунінгу перед навчанням для відеотрансформерів.

Другий розділ містить опис оптимальних умов для аналізу сцен, таких як використання чотирьох планів з одним ключовим кадром для визначення переходів та контексту. Детально описано проблеми, пов'язані зі збільшенням обсягу даних для аналізу, такі як зростання моделі та перенавчання, а також труднощі з короткими сценами, характерними для рекламного контенту, та запропоновано шляхи їх вирішення. Продемонстровано, як розроблений підхід розбиття відео на плани та

сцени значно скорочує кількість викликів моделей для визначення атрибутів, зберігаючи високу точність. Обґрунтовано потенціал подальшої оптимізації, зокрема методом прунінгу перед навчанням візуальних трансформерів.

Третій розділ містить опис покращеного методу прунінгу, що демонструє високу ефективність для оптимізації візуальних трансформерів для відео, значно зменшуючи обчислювальні витрати при збереженні точності. Описано дослідження підходів до прунінгу трансформерів, що призвело до модифікованого методу, який включає покращене визначення важливості ваг у SNIP через врахування активацій уваги та інтеграцію розрідженості матриць 2:4. Наведено експериментальні дані, що показують ~10% пришвидшення на неспеціалізованому обладнанні з потенціалом до більшого пришвидшення на спеціалізованих GPU.

Четвертий розділ містить опис розробленої архітектури програмного забезпечення, що ефективно використовує розбиття відео на сцени та плани для значного пришвидшення визначення відеоатрибутів. Представлено архітектурні рішення, які дозволяють гнучке використання системи на окремих машинах або в хмарних середовищах, із підтримкою Docker для інтеграції нових методів.

Дисертаційна робота оформлена відповідно до вимог наказу МОН України від 12 січня 2017 р. №40 «Про затвердження вимог до оформлення дисертації».

Оприлюднення результатів дисертаційної роботи

Наукові результати дисертації висвітлені у наукових публікаціях здобувача, серед яких: 4 статті у фахових наукових виданнях України. Із них одна наукова стаття опублікована у періодичному виданні, що входить до наукометричної бази Scopus та відноситься до першого — третього кuartилів (Q1—Q3) відповідно до класифікації SCImago Journal and Country Rank або Journal Citation Reports.

Також результати дисертації були апробовані на 6 наукових фахових конференціях.

Усі публікації здобувача мають належний науковий рівень, в них достатньо повно описані головні наукові здобутки, представлені в дисертації. Особистий внесок здобувача до всіх наукових публікацій, опублікованих у співавторстві, є вагомим.

Таким чином, наукові результати, описані в дисертаційній роботі, повністю висвітлені у наукових публікаціях здобувача.

Недоліки та зауваження до дисертаційної роботи

1. У першому розділі при аналізі вимог до програмного забезпечення (Таблиці 1.1 та 1.2) варто було б надати більш детальне обґрунтування деяких нефункціональних вимог, зокрема, щодо критеріїв надійності та продуктивності системи в контексті специфіки аналізу відео.

2. У другому розділі, де описується метод розбиття відео на сцени за допомогою візуального трансформера, було б доцільно навести порівняльний аналіз із більшою кількістю сучасних підходів, зокрема тих, що використовують інші архітектури нейронних мереж або їх комбінації для аналогічних завдань.

3. У третьому розділі, присвяченому оптимізації моделей, хоча й наведено результати експериментальної оцінки розробленого методу прунінгу, було б корисним розширити аналіз впливу різних рівнів прунінгу на точність та швидкодію для різних типів відеоконтенту.

4. У четвертому розділі, варто було б детальніше зупинитися на питаннях масштабованості розробленої розподіленої архітектури та навести кількісні показники її роботи при різному навантаженні та конфігураціях обчислювальних ресурсів.

Вважаю, що висловлені зауваження не є визначальними, не зменшують загальну наукову новизну та практичну значимість отриманих результатів та не впливають на позитивну оцінку дисертаційної роботи.

Висновок про дисертаційну роботу

Вважаю, що дисертаційна робота здобувача ступеня доктора філософії Здора Костянтина Андрійовича на тему «Моделі та програмні засоби підвищення швидкодії визначення відеоатрибутів за допомогою розбиття на сцени» виконана на високому науковому рівні, не порушує принципів академічної доброчесності та є закінченим науковим дослідженням, сукупність теоретичних та практичних результатів якого розв'язує актуальне наукове завдання, що має істотне значення для галузі знань 12 – Інформаційні технології.

Дисертаційна робота за актуальністю, практичною цінністю та науковою новизною повністю відповідає вимогам чинного законодавства України, що передбачені в п.6 - 9 «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії», затвердженого Постановою Кабінету Міністрів України від 12 січня 2022 р. №44.

Здобувач Здор Костянтин Андрійович заслуговує на присудження ступеня доктора філософії в галузі знань 12 - Інформаційні технології за спеціальністю 121 – Інженерія програмного забезпечення.

Офіційний опонент:

Завідувач кафедри Інформаційних систем
та технологій Державного університету
інформаційно-комунікаційних технологій,
доктор технічних наук, професор

Тетяна

Каміла СТОРЧАК

16.06.2025

